

**EL ORDENADOR AL SERVICIO DEL APRENDIZAJE DE  
LENGUAS: GRAMÁTICA Y DICCIONARIO  
AUTOMÁTICOS DEL ESPAÑOL\***

María Concepción Álvarez LEBREDO (Universidade de Santiago de  
Compostela)

**Abstract**

*In second language learning, it is essential to know the syntactical-semantic valency of verbs. Non-native speakers may take advantage of tools, such as grammars and dictionaries. However, in the case of Spanish, those tools, although they give a great deal of data, are not suitable for L2 acquisition. The valency information they contain is not sufficiently explicit to allow the encoding of admissible messages in Spanish. Here we try to show an automatic recognition system which first analyses the verb constructions proposed by the user and secondly shows the verb meaning which is implied in the example. Grammar, dictionary and machine joint efforts in the process of learning.*

**Key-words:** *valency; formal grammar; electronic dictionary; interactive learning.*

**Resumo**

*Na aprendizagem de uma segunda língua, é essencial o conhecimento da valência sintático-semântica dos verbos. As gramáticas e os dicionários são algumas das ferramentas mais importantes de que se servem os falantes não-nativos. Mas, no caso do espanhol, essas ferramentas, embora forneçam muitas informações, geralmente não são adequadas para a aquisição da língua, porque a informação valencial não é suficientemente explícita para que o falante não-nativo possa codificar mensagens admissíveis em espanhol. Apresentamos aqui um sistema de reconhecimento automático que analisa as construções verbais*

---

\* *the ESPecialist* agradece a Claudia Jacobi pela revisão do idioma espanhol

*propostas pelo usuário e indica a aceção do verbo implicada no exemplo. Gramática, dicionário e máquina aliam-se no processo de aprendizagem.*

**Palavras chave:** *valência; gramática formal; dicionário eletrônico; aprendizagem interativa.*

## 1. Valencia verbal y aprendizaje de la lengua

Es bien conocida la teoría de la valencia según la cual cada verbo condiciona la aparición de un número determinado de “casillas” o “lugares vacíos” que llamamos *actantes* o argumentos<sup>1</sup>. Así, el verbo *atribuir* requiere la presencia de tres argumentos para constituir una secuencia gramatical en español:

*alguien - atribuye - algo - a alguien*

Para poder cifrar correctamente en español un mensaje que contenga un verbo, como por ejemplo **atribuir**, necesitamos la siguiente información:

- el número de actantes con que se construye o *valencia cuantitativa*. En el ejemplo, tres;
- las funciones sintácticas que desempeñan esos actantes (*valencia cualitativa*). En el caso del verbo **atribuir**, las funciones implicadas son **sujeto**, **complemento directo** y **complemento preposicional**;
- el carácter obligatorio o facultativo de los actantes. En el ejemplo, son todos obligatorios excepto el **sujeto**, que puede ser omitido:

---

<sup>1</sup> Fue Lucien Tesnière quien introdujo el concepto de valencia en los estudios lingüísticos. A continuación, autores como Busse (1974), Helbig y Schenkel (1991), Viléla (1992), Sommerfeldt (1983 a e b) o Allerton (1982) aplicaron dicho concepto a la descripción de diferentes lenguas. Sus obras constituyen puntos de referencia obligados para el estudio de la teoría de la valencia. Cabe destacar la aplicación a la lengua portuguesa que hacen Busse y Viléla (1986), el tratamiento de Somers (1987) desde el punto de vista de la lingüística computacional o la confección de diccionarios bilingües como el de Dietrich (1980).

**atribuye** - *algo* - *a alguien*

Pero:

*\*alguien atribuye algo*

*\*alguien atribuye a alguien*

*\*alguien atribuye*

- los tipos de unidad o categorías gramaticales que pueden desempeñar las funciones. En la teoría de la valencia, a esta información se le denomina *rasgos de subcategorización*. En el ejemplo, **sujeto** y **c.directo** son desempeñados por frases nominales, mientras que la función **c.preposicional** la realiza una frase preposicional encabezada por la preposición *a*. La información sobre la preposición regida es esencial para la correcta utilización del verbo;
- las *restricciones de selección* que el verbo impone sobre sus argumentos, esto es, los rasgos semánticos que han de tener los elementos que acompañan al verbo dado. Por ejemplo, el verbo **atribuir** exige un **sujeto** con el rasgo [+animado].

*alguien - atribuye - algo - a alguien*

*\*algo - atribuye - algo - a alguien*

Conociendo toda esta información, el hablante debe ser capaz de codificar adecuadamente su mensaje:

*Los críticos atribuyen el cuadro a Sorolla*

*Atribuyen el cuadro a Sorolla.*

Toda esa información pertenece a la valencia o régimen del verbo. Es información predecible, independiente del contexto. Constituye el conocimiento básico que necesita el hablante para usar un verbo correctamente. Muchos de los errores cometidos por hablantes no nativos de una lengua se deben al desconocimiento

parcial o total de la información valencial asociada al verbo en la lengua en cuestión<sup>2</sup>. Es frecuente que el hablante transfiera la información valencial de su propia lengua cuando intenta expresarse en otra. Un caso sencillo de error de estas características es la confusión de la preposición regida. Así, un francés podría decir erróneamente *\*alguien piensa a algo*, en lugar de *alguien piensa en algo*, porque traslada la preposición que rige el verbo *penser* (à) en francés al verbo español *pensar*.

## 2. Gramática y diccionario

Los instrumentos a los que recurren los hablantes a los que se les plantean dudas sobre una determinada construcción, son los diccionarios y, en menor medida, las gramáticas. Está demostrado que se acude antes a un diccionario que a una gramática o cualquier otra obra de consulta. El diccionario es la herramienta preferida de los usuarios, lo cual es perfectamente comprensible dada la facilidad y rapidez con que se puede efectuar la consulta. Pero, a menudo, no se encuentra en ellos la información que se precisa para construir una secuencia correcta. Pueden servir para conocer el significado de un verbo o palabra concreta, siempre que seamos capaces de discernir cuál de las acepciones, de entre las que recoge el diccionario, es la que se adecua al ejemplo que se está intentando descifrar. Más difícil le será construir un mensaje que contenga un verbo en una acepción específica, utilizando exclusivamente los datos que el diccionario proporciona.

Durante mucho tiempo, la mayor parte de los diccionarios limitaban sus referencias a la valencia sintáctica del verbo a abreviaturas como *tr.* (transitivo), *pron.* (pronominal), etc. En raras ocasiones, incluían ejemplos e información sobre preposición regida. El magnífico *Diccionario de régimen y construcción* de Cuervo suministraba abundante información sobre régimen verbal, pero no resultaba adecuado para el aprendizaje de una lengua,

<sup>2</sup>La aplicación de la teoría de la valencia a la enseñanza de lenguas ha sido tratada, entre otros, por Viléla (1992) y por Fink (1977).

debido, entre otras cosas, a la complejidad y extensión de las entradas y a las referencias a estados anteriores de nuestra lengua, como el Español Clásico.

En la actualidad, existen numerosos diccionarios que contienen una valiosa información explícita e implícita sobre aspectos relacionados con la valencia sintáctico-semántica del verbo<sup>3</sup>, entre ellos, el *Diccionario de Uso del Español* de María Moliner (DUE), el *Salamanca*, o el *Clave*, por citar unos pocos. El *Salamanca* marca en la definición el elemento que funciona como sujeto encerrándolo entre ángulos y los que funcionan como objetos, encerrándolos entre corchetes, siempre con indicación de si se trata de una persona o una cosa, entre otras especificaciones semánticas. El DUE señala con flechas oblicuas los elementos de la definición que pueden funcionar como argumentos del verbo definido. El *Clave* utiliza la paráfrasis *referido a...*, para marcar las restricciones semánticas o ámbito de aplicación del verbo. Todos ellos recurren a los ejemplos, de los que se puede extraer ciertas conclusiones acerca del régimen.

Los diccionarios de este tipo suponen un gran avance en el terreno lexicográfico. Con todo, en la mayoría de los casos, este tipo de herramientas no satisface las necesidades del usuario que espera encontrar indicaciones claras que le permitan usar un verbo u otra clase de palabra, de forma espontánea, sin tener que contentarse con repetir el ejemplo que se propone en el diccionario. Es necesario un diccionario en que aparezca recogida toda la información relevante para cifrar mensajes<sup>4</sup>. Para cubrir esa necesidad proponemos aquí un método alternativo.

---

<sup>3</sup> En la tesis de García Martínez (1996) se da un repaso a la información semántica contenida en los principales diccionarios del español..

<sup>4</sup>El modelo de diccionario que proponemos para el español es semejante al diccionario inglés *COLLINS COBUILD* que proporciona en las definiciones suficiente información para emplear las palabras en un contexto apropiado.

### 3. Gramáticas automáticas - el formalismo AGFL

Las gramáticas describen las normas y principios que rigen el sistema de una lengua. Habitualmente son obras de difícil consulta. No siempre es fácil encontrar en ellas la información que buscamos. Por otra parte, en una gramática no podemos encontrar la información relativa al uso de piezas léxicas concretas, por lo menos no de todas ellas; eso es una labor que le corresponde al diccionario.

Las gramáticas automáticas son conjuntos de reglas escritas en un formalismo comprensible por el ordenador. Frente a las gramáticas tradicionales, escritas en papel, el usuario de una gramática automática o formal no necesita consultar o buscar en ella la teoría sobre un determinado fenómeno o aspecto de la lengua. La podrá utilizar para analizar ejemplos que él mismo proponga. Un programa llamado *generador de analizadores* crea, a partir de la gramática formal que escribe el lingüista, un *parser* o analizador automático. Gracias a este analizador, el usuario podrá proponer cadenas lingüísticas que el ordenador analizará en fracciones de segundo.

La gramática automática que constituye la propuesta de este artículo está escrita en el formalismo *AGFL*<sup>5</sup> (de *Affix Grammars Over a Finite Lattice*) desarrollado en la Universidad de Nimega por el profesor C.H.A. Koster. Se trata de un formalismo especialmente diseñado para el tratamiento de lenguajes naturales. Es un formalismo de dos niveles en que las reglas de reescritura son incrementadas por medio de parámetros o *afijos* que permiten establecer relaciones entre los distintos elementos de la gramática.

El *parser* resultante de la compilación de la gramática *AGFL* permite analizar sintáctica (y/o morfológicamente) cadenas lingüísticas introducidas por el usuario (*input*) y ofrece el árbol de

---

<sup>5</sup>Disponible en la red en la dirección: <http://www.cs.kun.nl/agfl>

derivación o *parse tree* correspondiente donde podemos observar las relaciones gramaticales que se establecen entre los elementos de la cadena. Al mismo tiempo, funciona como reconocedor automático, ya que sólo analizará secuencias gramaticales, esto es, secuencias que se ajusten a las reglas que constituyen la gramática. Aquellas cadenas que no son admisibles de acuerdo con la gramática formal, no serán analizadas.

El *output* de la secuencia *ellos comen patatas* podría tener una forma como la siguiente, utilizando las convenciones del formalismo *AGFL*:

(1) *ellos comen patatas*

cláusula

SUJETO (tercera, plural, masculino)

pron personal (nominativo, tercera, plural, masculino)

“ellos”

PREDICADO (su cd, tercera, plural)

verbo (su cd, tercera, plural)

“comen”

CDIRECTO (tercera, plural, femenino)

sustantivo (comun, tercera, plural, femenino)

“patatas”

parsing 1

all : 1<sup>6</sup>

En cambio, la secuencia *\*ellos come patatas* no sería analizada porque rompería una regla de la gramática que establece que el **sujeto** y el **predicado** han de concordar en número y persona.

(2) *\*ellos come patatas*:

---

<sup>6</sup> La indicación all: 1 hace referencia al hecho de que la gramática sólo encuentra un análisis posible para la secuencia dada. Si la secuencia fuese ambigua mostraría tantos análisis como segmentaciones o interpretaciones funcionales distintas permitiera la gramática: dos, tres o incluso más. Del mismo modo, la indicación all: 0 indica que el *parser* no encuentra ningún análisis adecuado a las reglas de la gramática formal.

>> ellos come patatas (INPUT)

all : 0 (OUTPUT)

En el ejemplo (1), el analizador sintáctico ha reconocido la secuencia que proponía el usuario como perteneciente a la lengua descrita por la gramática y la ha etiquetado como cláusula. En el árbol de derivación de más arriba se puede ver que el *parser* ha descompuesto la secuencia *ellos comen patatas* en sus constituyentes funcionales. En la notación que utilizamos, representamos las funciones sintácticas con letras mayúsculas y las categorías gramaticales con minúsculas. En el seno de la secuencia del ejemplo se establecen tres relaciones funcionales: la función **sujeto** está realizada por la categoría gramatical *pronombre*; el **c.directo**, por un *sustantivo*; mientras que un verbo desempeña la función que llamamos **predicado**.

El modelo sintáctico subyacente en esta representación es de tipo constitutivo-funcional. Considera las funciones y las categorías como primitivos de la teoría. Se trata de un modelo de análisis no binario que descompone las unidades de la lengua en sus constituyentes funcionales y estos en los suyos, y así sucesivamente hasta llevar al nivel de la palabra<sup>7</sup>. En el ejemplo sólo hay un nivel de segmentación porque los elementos funcionales de esta cláusula son los elementos mínimos (palabras); pero si se tratase de unidades más complejas obtendríamos dos, tres e incluso más niveles de análisis.

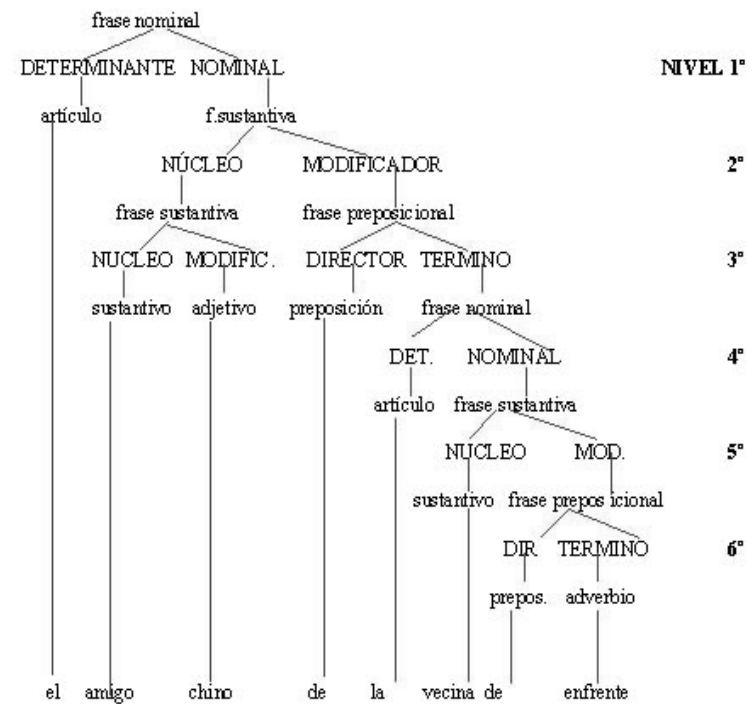
(3) *El amigo chino de la vecina de enfrente.*

En el ejemplo (3) se pueden distinguir 6 niveles de análisis:

---

<sup>7</sup>Este modelo teórico se describe en Rojo y Juliá (1989).





Este análisis podría tener en EGFL la siguiente salida:

>> el amigo chino de la vecina de enfrente

frase nominal

DET (masculino, singular)

artículo(masculino, singular)

“el”

NOMINAL (masculino, singular)

frase sustantiva (masculino, singular)

NUCLEO (masculino, singular)

frase sustantiva (masculino, singular)

NUCLEO (masculino, singular)

sustantivo (masculino, singular)

“amigo”

MODIFICADOR (masculino, singular)

adjetivo (masculino, singular)

“chino”

MODIFICADOR

frase prep (de)

DIR (de)

prep (de)

“de”

TERMINO

frase nominal (femenino, singular)

DET (femenino, singular)

artículo (femenino, singular)

“la”

NOMINAL (femenino, singular)

frase sustantiva (femenino, singular)

NUCLEO (femenino, singular)

sustantivo (femenino, singular )

“vecina”

MODIFICADOR

frase prep (de)

DIR (de )

prep (de)

“de”

TERMINO

adverbio

“enfrente”

El formalismo *AGFL* ha sido utilizado en diferentes aplicaciones, como por ejemplo, la recuperación automática de información (*information retrieval*). Proponemos aquí utilizarlo también para el aprendizaje de lenguas.

#### **4. El ordenador aplicado al aprendizaje: una gramática y un diccionario complementarios**

Hemos visto más arriba las características más relevantes del formalismo *AGFL*. También hemos visto el aspecto que muestra la salida del análisis realizado por el *parser*. En esta salida,

aparecen un conjunto de *cadena no terminal* (por ej.: **sujeto, verbo, prep, c.directo...**) acompañadas de una información encerrada entre paréntesis y separada por comas: los *afijos*. Como hemos dicho, los afijos permiten establecer relaciones entre elementos dentro de la gramática. Son pues, requisitos del formalismo y del sistema. Son mecanismos necesarios para su correcto funcionamiento; tienen un fin instrumental, no tienen valor en sí mismos. Pero, secundariamente, ofrecen una información al usuario del sistema. El *no terminal*

PREDICADO ( su cd, tercera, plural )

que aparece en el ejemplo (1), *ellos comen patatas*, contiene tres argumentos o afijos. El sentido del afijo su cd es garantizar la combinación correcta del **predicado** con los demás constituyentes funcionales. En este caso, el **predicado** realizado por el verbo *comen* se construye con **sujeto** (su) y **c.directo** (cd). Los afijos tercera y plural sirven para asegurar la concordancia entre **sujeto** y **predicado**. Para que sea reconocida la unidad cláusula tiene que haber coincidencia entre los valores de los afijos **numero** y **persona** del **sujeto** y del **predicado**. Se dice entonces que los valores *unifican*. Sin ello no es posible el reconocimiento. Pero al mismo tiempo, el usuario del sistema obtiene una información interesante. De igual manera, en

pron personal (nominativo, tercera, plural, masculino)  
“ellos”

de un solo golpe de vista, el usuario conoce la categoría y los rasgos que definen al terminal “ellos”. Se trata de la tercera persona plural de pronombre personal, en caso nominativo. La información en sí misma es irrelevante para el sistema y su funcionamiento. Pero constituye una información valiosa para el estudioso de la lengua que contempla el resultado del análisis.

Nuestra propuesta consiste en utilizar esa información como fin, no como medio; convertirla en aspecto esencial de la gramática: ya que en el *output* no se ocultan los afijos, los emplearemos como medio fundamental de comunicación con el usuario.

El objeto de nuestros esfuerzos es el verbo. Queremos dotarlo de toda la información valencial que le corresponde, de todos los rasgos que lo definen y que permiten al usuario diferenciarlo de otros verbos homónimos o de otras acepciones del mismo verbo y que, al mismo tiempo, permiten crear mensajes correctos en los que aparezca el verbo dado. La gramática *AGFL* resultante está articulada de tal modo que el análisis de una construcción verbal arrojado por el *parser* nos mostrará como afijos, es decir, encerrada entre paréntesis, toda la información valencial necesaria para el estudioso de la lengua. Esa información consiste, como hemos dicho más arriba, en: los elementos funcionales con los que se puede combinar, los rasgos de subcategorización de esos elementos y las restricciones semánticas que impone sobre ellos.

Veamos algunos ejemplos sencillos de acepciones de verbos que implican un solo actante.

Consideremos, por ejemplo, el verbo **quejar** en su acepción

“Proferir gritos o sonidos con que se muestra dolor físico o padecimiento moral” (DUE)<sup>8</sup>

Nuestra gramática reconoce la secuencia

(4) *alguien se queja*:

>> alguien se queja

---

<sup>8</sup> Utilizamos las definiciones del diccionario de María Moliner por parecernos uno de los más completos y coherentes en los aspectos que estamos considerando.

frase verbal (indep, uno, ind)  
 ARG (SUJ, an, terc, sing)  
 fn (an, terc, sing)  
 PRED ( uno, su pro, anim, terc, sing, ind )  
 verbo (uno, su pro, anim, terc, sing, ind)  
 “se”  
 “queja”

parsing 1  
 all: 1

Pero no analizará

(5) *alguien queja:*

>> alguien queja  
 all: 0

ya que el verbo **quejar** exige, y así está establecido en nuestra gramática formal, la presencia obligatoria del clítico.

Tampoco analizará

(6) *alguien me queja:*

>> alguien me queja  
 all: 0

En este caso, el motivo es que la secuencia propuesta por el usuario viola una de las reglas de la gramática según la cual el clítico de los verbos pronominales ha de concordar en número y persona con la forma verbal a que acompaña.

Una consulta con la secuencia

(7) *algo se queja,*

>>algo se queja

all: 0

supondrá un nuevo fracaso porque el ejemplo no se ajusta a las restricciones de selección que el verbo impone a su **sujeto**: ha de ser animado. Pero el pronombre indefinido *algo* corresponde a un referente inanimado.

Volvamos al ejemplo (4) - *alguien se queja* - y fijémonos en el *no terminal* que define al verbo:

*verbo (uno, su pro, anim, terc, sing, ind)*

Indica todas las características de que hemos hablado:

- valencia cuantitativa: uno. Nos indica que es un verbo monovalente; se construye con un solo actante;
- la clase a la que pertenece: su pro. Quiere decir que el actante que lo acompaña ha de desempeñar la función de **sujeto** (su) y ha de ir acompañado de un pronombre clítico (pro),
- los rasgos del actante: ha de ser una frase nominal animada. Ese es el significado del afijo *anim*.

Los tres afijos restantes proporcionan información morfológica: tercera persona, número singular, modo indicativo.

Los tres primeros afijos son los que contienen la información valencial. Son los que necesita el usuario para codificar de forma correcta su mensaje. Como veremos, esos tres afijos también servirán para diferenciar el significado del verbo en ese contexto frente al significado de otros verbos homófonos o de otras acepciones del mismo verbo.

Veamos otro ejemplo:

(8) *alguien anda*

>> alguien anda

frase verbal (indep, uno, ind)

ARG (SUJ, an, terc, sing)

fn (an, terc, sing)

“alguien”

PRED (uno, su, anim, terc, sing)

verbo (uno, su, anim, terc, sing)

“anda”

Por una parte, la gramática sólo reconoce las construcciones posibles. Por otra, individualiza la acepción del verbo a que el ejemplo hace referencia, mostrándonos sus características valenciales. Hay que destacar que la definición que se nos muestra en el árbol de derivación, es la que corresponde exclusivamente al ejemplo que el usuario ha puesto como *input*. En el ejemplo, no se hacen referencias a otras acepciones del verbo **andar** que implican **sujeto** inanimado, sino tan solo a la definición

“Caminar. Andar. Moverse de una lado para otro dando pasos” (DUE)

Si el *input* hubiese sido

(9) *algo anda*,

entonces obtendríamos el siguiente análisis

>> algo anda

frase verbal (uno, indep, ind)

ARG (SUJ, in, terc, sing)

fn (in, terc, sing)

“algo”

PRED (uno, su, inan, terc, sing, ind)

verbo (uno, su, inan, terc, sing, ind)

“anda”

donde la definición del verbo ha variado:

verbo (uno, su, inan, terc, sing, ind)

Corresponde a otra acepción del mismo verbo.

El aspecto más innovador de esta gramática escrita en el formalismo *AGFL* es su conexión con el diccionario. Los *no terminales* que definen los verbos se corresponden con las entradas de un diccionario electrónico complementario de la gramática.

El *no terminal* correspondiente al **verbo** de (8) *-alguien anda-* se refiere a la siguiente entrada de nuestro diccionario electrónico:

verbo(uno,su,anim): Alguien anda si camina, si se mueve de un lado para otros dando pasos.

La entrada del diccionario recoge los tres afijos más significativos de los que aparecen en la representación resultante del proceso de *parsing*. Son los tres afijos que corresponden a la información valencial del verbo.

El usuario que introduce como *input* el ejemplo (9) *-“algo anda”-* obtiene, además del análisis sintáctico con la forma que hemos reproducido más arriba, la siguiente definición del diccionario:

Algo anda si se mueve mecánicamente cambiando de lugar.



Un aspecto interesante es que cuando en el contexto no hay suficiente información para individualizar la única acepción del verbo, el sistema nos muestra en el análisis sintáctico todos los rasgos posibles en ese contexto determinado. Y, al mismo tiempo, todas las definiciones que podrían corresponder al verbo en cuestión:

(10)

>> anda

frase verbal (indep, uno, ind)

PRED (uno, su, anim|inan, terc, sing, ind)

verbo (uno, su, anim|inan, terc, sing, ind)

“anda”

1. Alguien anda si camina, si se mueve de un lado para otro dando pasos.
2. Algo camina si se mueve mecánicamente cambiando de lugar.

Al no estar explícito el **sujeto** en (10), el resultado del análisis recoge las dos posibilidades: anim|inan. Con ello remite a dos acepciones distintas, a dos definiciones: la primera implica un **sujeto** animado; la segunda, inanimado.

La utilidad de este sistema es doble. Por una parte, el usuario obtiene exactamente la definición adecuada al ejemplo que le preocupa. Todos sabemos lo desesperante que resulta buscar en un diccionario un verbo con la intención de aclarar una construcción cuyo significado desconocemos y encontrarnos con una columna repleta de acepciones y sentidos distintos. En estos casos, tenemos que desplegar una actividad casi detectivesca para descubrir cuál de ellas es la que conviene a nuestro ejemplo. A menudo, tenemos que aplicar mentalmente una a una todas las definiciones a nuestro contexto, para intentar deducir cuál tiene sentido. Otra ayuda son los ejemplos, cuando los hay. En este caso, lo habitual es prestar atención a aquellos que nos parecen más semejantes al nuestro. En el sistema que estamos presentando, no son necesarios estos esfuerzos puesto que, como hemos visto, el sistema selecciona por

nosotros la acepción o acepciones convenientes en cada caso. Una ventaja clara del sistema es que no hay necesidad de interpretar las abreviaturas y símbolos del diccionario para dar con la definición que nos interesa.

Por otra parte, el usuario puede profundizar un poco más en la gramática para intentar aprender con un pequeño esfuerzo aspectos relacionados con las construcciones verbales del español. Si el analizador rechaza la construcción producida por el usuario, o bien la acepta con una definición inadecuada, muy diferente a la que esperábamos, el usuario-aprendiz de la lengua, guiándose por el *no terminal* correspondiente al verbo que se muestra en la representación arbórea, puede modificar su propuesta de *input* hasta dar con la construcción correcta.

Supongamos que alguien equivocadamente cree que el verbo **arrojar** puede construirse sólo con **sujeto** animado y sin **c.directo** con el significado de ‘lanzar, sacar (la pelota)’ basándose, por ejemplo, en un diccionario bilingüe que no diferencia con claridad las acepciones de los verbos y no contiene ejemplos. Si intenta analizar la construcción *alguien arroja*, obtendrá el siguiente resultado:

frase verbal (uno, indep, ind)

ARG (SUJ, an, terc, sing)

fn (an, terc, sing)

“alguien”

PRED ( uno, su, anim, terc, sing, ind)

verbo (uno, su, anim, terc, sing, ind)

“arroja”

que nos remite a la entrada del verbo (uno, su, anim) del diccionario electrónico, la cual lleva aparejada la siguiente definición:

(vulgar) Alguien arroja cuando vomita, cuando expulsa por la boca el contenido del estómago.

El usuario comprueba claramente que la definición no corresponde a su idea inicial y deberá probar nuevas construcciones que la gramática admita y cuya definición le parezca adecuada como “alguien arroja algo”; o bien, probar con nuevos verbos: “alguien lanza”; “alguien lanza algo”...

Se convierte así en un sistema hasta cierto punto interactivo, que requiere algún esfuerzo por parte del usuario. Este esfuerzo convierte el aprendizaje en significativo. Es un factor que no está presente en los diccionarios y gramáticas convencionales.

## 5. Conclusión

Se presenta aquí una gramática formal con un diccionario electrónico asociado que inciden sobre el aspecto de codificación del lenguaje, frente al de mera descodificación de la mayoría de los sistemas de este tipo. Otras gramáticas formales existentes se contentan con poder analizar secuencias, con reconocer todas las secuencias posibles de una lengua. Nuestro sistema se propone un objetivo más ambicioso: reconocer las construcciones verbales posibles y rechazar las imposibles. Sólo así se puede aplicar al aprendizaje de una segunda lengua.

Recebido em: 06/1997. Aceito em: 11/1997.

## Referencias

- ALLERTON, D.J. (1982) *Valency and the English verb*. London: Academic Press.
- BUSSE, W. (1974) *Klasse, transitivität, valenz. Transitive klassen des verbs in französischen*. München: Fink Verlag.
- BUSSE, W. & M. VILÉLA (1986) *Gramática de valências: apresentação e esboço de aplicação à língua portuguesa*. Coimbra: Livraria Almedina.

- CLAVE - *Diccionario de uso del español actual* (1996) Madrid: Ediciones SM.
- Collins Cobuild. *English language dictionary* (1987) London: Collins.
- CUERVO, R.J. (1994) *Diccionario de construcción y régimen de la lengua española* (continuado y editado por el Instituto Caro y Cuervo). Santafé de Bogotá: Instituto Caro y Cuervo.
- Diccionario Salamanca de la lengua española* (1996) Santillana: Universidad de Salamanca.
- DIETRICH, R. (1980) *Diccionario de valencias verbales alemán-español*. Tübingen: Gunter Narr.
- FINK, S.R. (1977) *Aspects of a pedagogical grammar based on case grammar and valency theory*. Tübingen: Niemeyer Verlag.
- HELBIG, G. & W. SCHENKEL (1991) *Wörterbuch zur valenz und distribution deutscher verben*. Tübingen: Max Niemeyer Verlag.
- KOSTER, C.H.A. (1991) Affix grammars for natural languages. IN: *Attribute grammars, applications and problems. International Summer School, SAGA, Praga. Lecture Notes in Computer Science, 545*. Berlin: Springer-Verlag.  
Disponible en <http://www.cd.kun.nl/agfl/Papers.html>.
- MARTÍNEZ, G. J. (1996) *Gramática y diccionario. El prefijo RE-*. Tesis doctoral. Madrid: Universidad Autónoma de Madrid.
- MOLINER, M. (1994) *Diccionario de uso del español*. Madrid: Gredos.
- ROJO, G. & T. J. JULIÁ (1992) *Fundamentos del análisis sintáctico funcional*. Lalia 2, Universidad de Santiago de Compostela.
- SOMMERFELDT, K-E (1983a) *Wörterbuch zur valenz und distribution der substantive*. Tübingen: Max Niemeyer.
- \_\_\_\_\_ (1983b) *Wörterbuch zur valenz und distribution deutscher adjektive*. Tübingen: Max Niemeyer.
- SOMERS, H.L. (1987) *Valency and case in computational linguistics*. Edimburgh: Edinburgh University Press.

- TESNIÈRE, L. (1969) *Éléments de syntaxe structurale*. París: Klincksieck.
- VILÉLA, M. (1992) A gramática de valências como ‘base’ para o ensino das línguas. IN: M. VILÉLA. *Gramática de valências. Teoria e aplicação*. Coimbra: Livraria Almedina.

*María Concepción Álvarez Lebrede belongs to the Department of Spanish Filology, Theory of Literature and General Linguistics of the University of Santiago de Compostela, where she has a research studentship. She belongs to the Spanish Syntax Group (<http://www.usc.es/~sintx>), with which she has collaborated in several projects related to the structure of the clause and the automatic analysis of Spanish verb government. Nowadays, she works in the DoRo project (financed by the UE) and she is working on her Ph.D. entitled "Automatic Recognition of Spanish Verb Constructions". Her areas of interest are: Spanish syntax, particularly formal syntax; corpus linguistics; computational linguistics applied to Spanish learning, and ‘parsers’.*

*e-mail: femcal@uscmail.usc.es*